

Is Mathematical Rigor Necessary in Physics?

Kevin Davey

September 4, 2002

Abstract

Many arguments found in the physics literature involve concepts that are not well-defined by the usual standards of mathematics. I argue that physicists are entitled to employ such concepts without rigorously defining them so long as they restrict the sorts of mathematical arguments in which these concepts are involved. Restrictions of this sort allow the physicist to ignore calculations involving these concepts that might lead to contradictory results. I argue that such restrictions need not be *ad-hoc*, but can sometimes be justified by considering some of the metaphysical issues surrounding the question of the applicability of mathematics to physical reality.

1 Introduction.

Physics textbooks are full of arguments that are unpersuasive insofar as we try to construe them as pieces of mathematics.¹ Often, the reason these arguments are unpersuasive as pieces of mathematics is that they violate the commonly accepted standards of mathematical rigor. Some such violations are fairly innocent, and need not bother anyone too greatly – as, for instance, when limit signs or derivatives are interchanged without the scrutiny ordinarily required by the mathematician. But other violations are not obviously so innocent. In particular, the introduction and use of mathematically ill-defined concepts – as occurs, for instance, with the use of delta functions and path integrals in quantum mechanics – cannot be obviously dismissed as merely insignificant violations of the usual standards of mathematical rigor. Such violations instead raise what can sometimes be difficult questions for those interested in a rigorous reconstruction of the arguments of physics.

The remarkable thing, however, is that in spite of the difficult questions that such violations raise, mathematically unrigorous arguments can – at least, when things go well –

¹In order for an argument to be persuasive as a piece of mathematics, I do not require that it be written out as a proof in ZFC, or even that it be rigorous by the most exacting of mathematical standards. By calling a piece of reasoning persuasive as a piece of mathematics, I simply mean that it is the sort of thing that most mathematicians would be inclined to accept as a valid piece of mathematical reasoning.

settle questions of physics with great decisiveness, obliterating opposing schools of thought. It is strange that the standards of rigor and exactness lying at the heart of the mathematical method should be so dispensable in this way when we try to apply mathematics to the physical world. Some good examples of this are provided by Steiner in [23], where he invites philosophers to grapple with the broader epistemological puzzles that this phenomenon raises. In this paper, I wish to accept Steiner's challenge, and assess what our attitude to all of this should be.

There are at least three views that might be taken on the question of what role the mathematicians' standards of rigor ought to play in physics. According to the first view, the physicist behaves irresponsibly when he formulates and allows himself to be moved by mathematically unpersuasive arguments, and should be criticized accordingly. According to this view, the physicist ought only to be moved by arguments that are mathematically rigorous. I will call this the *conservative* view of the role of mathematical rigor in physics.

According to the second view, until an area of physics reaches full maturity, physicists working in that area will likely be forced to traffic in arguments that are not fully mathematically rigorous. Given this fact, it would be unfair to criticize a physicist working in an immature area of physics for being occasionally moved by mathematically unrigorous arguments. According to this view, however, once an area of physics reaches full maturity, we must demand that the physicist working in it *only* be moved by mathematically persuasive arguments. I will call this the *moderate* view of the role of mathematical rigor in physics; the difference between it and the conservative view is that it allows for non-rigorous argumentation in immature physical theories.

As always, variations on these views are possible. For instance, even the conservative view might allow for 'innocent' violations of mathematical rigor, where a violation is deemed innocent if it is the sort of thing which it is reasonable to conjecture *could* be made rigorous in a straightforward way. As mentioned already, care-free interchanges of limit signs are generally regarded (rightly or wrongly) as innocent in this sense. Likewise, the moderate view could allow for such violations in the case of a mature theory. Such minor modifications of the conservative and moderate view may be adopted at the reader's discretion, and do not affect the subsequent argument.

In addition to the conservative and moderate views, there is a third view about the role of mathematical rigor in physics – I will call it the *liberal* view – according to which it is inappropriate to demand that even the most mature areas of physics traffic exclusively in mathematically rigorous concepts and arguments. According to this view, even in the most mature areas of physics, it is still sometimes appropriate to be persuaded by an argument

which is neither rigorous nor obviously rigorizable.²

Much mainstream philosophy of science rejects this liberal view outright. The positivist tradition, for instance, conceived of physics as a discipline in which all theoretical concepts could be (and indeed, *had* to be) rigorously defined, and in which theories were given by axiomatic systems in which all facts of interest to the physicist could be deduced as ‘theorems’ in a logically rigorous manner. An example of this is Carnap’s vision of science in the ‘Aufbau’ [4], in which he called for science to provide a logically rigorous ‘construction’ of its concepts from basic concepts. This is incompatible with the liberal view, which allows for the employment of concepts which do not meet – and perhaps cannot meet – the mathematician’s standards of well-definedness.³ Much subsequent philosophy of science also finds itself clinging to the conservative – or at best moderate – view of the role of mathematical rigor in physics. On p111 of [12], for instance, Hempel tells us that he views scientific theories, in their ‘advanced stages’, as being given by ‘deductively developed axiomatized systems’. In his influential essay ‘Studies in the Logic of Explanation’ [12], Hempel also argues that in any scientific explanation, the explanandum must be *logically deducible* from the explanans in order to be genuinely explanatory.⁴ This goes against the liberal view, which allows for the possibility of an explanation which cannot be made cogent as a piece of logic, insofar as it involves concepts which are not rigorously defined, and not obviously rigorously definable.⁵ Furthermore, even Popper argued that all physics should aim towards a ‘rigorous system’ – see in particular section 16 of [19], where Popper argues that this is important in order that we be able to decide unambiguously when a theory has been falsified by an observation.

Yet despite all this bad press, the liberal view enjoys great support amongst physicists, who often make disparaging remarks about the insistence on mathematical rigor, even in the most mature areas of physics. In their classic textbook [15], Landau and Lifshitz declare that ‘*no attempt has been made at mathematical rigor in the treatment, since this is*

²This, of course, is not to suggest that a rigorous physics is *impossible*, or never desirable.

³Note, however, that Carnap was clearly a moderate rather than a conservative on the role of standards of meaningfulness in the sciences. See pp288-289 of [4]: ‘*Thus, the formation of the constructional system is the first aim of science.* It is the first aim, not in a temporal, but in a logical sense. The historical development of science does not have to postpone the investigation of an object until this object is placed within a constructional system. For objects on higher levels, especially for biological and cultural objects, science must not wait for this to take place, if it does not want to forgo, for a long time, the development of these essential fields with their important practical applications.’

⁴See p247 of Hempel’s [12]: ‘The explanandum must be a logical consequence of the explanans; in other words, the explanandum must be logically deducible from the information contained in the explanans; for otherwise the explanans would not constitute adequate grounds for the explanandum.’

⁵In the main examples of nonrigorous mathematics I will be considering in this paper, one *can* conceive the relevant arguments and concepts as logically rigorous, so long as one is prepared to admit either inconsistent sets of premises or self-contradictory concepts. But given Hempel’s insistence that consistency be a basic requirement of the axioms and concepts of a scientific theory, the tension between Hempel and the liberal view remains.

anyhow illusory in theoretical physics ...'. In [16], Mandelbrot, replying to a comment by P. Anderson, says that '*...Anderson describes mathematical rigor as 'irrelevant and impossible'. I would soften the blow by calling it 'besides the point and usually distracting, even where possible.'*' In [2], Bridgman declares '*...the commitment of the physicist to the use of mathematics itself constitutes, paradoxically, a renunciation of the possibility of rigor.*' One also finds many well-read physicists citing Heaviside, who in reply to being criticized for the employment of unrigorous mathematical methods, rhetorically asked '*...should I refuse a good dinner simply because I do not understand the process of digestion?*' Finally, although the attitudes towards mathematics of physicists such as Feynman and Bohr are complicated enough to deserve a discussion on their own, it is clear that their views, although different from each other, were species of the liberal view.^{6 7}

In this paper, I wish to side with the physicists – and against Carnap et al. – by developing and defending the liberal view of the role of mathematical rigor in physics, according to which mathematically unrigorous arguments can sometimes be persuasive as pieces of physics. It must be emphasized that the liberal view does not assert that *any* violation of mathematical rigor ought to be permitted in physics, but only that *some* such transgressions ought to be. Consequently, in order to argue for the liberal view, it suffices to isolate just one important type of violation of mathematical rigor found in the physics texts, and defend its use. This is what I shall do in the present paper.

In the next section, I will identify the the type of violation that the present paper is to focus on – broadly speaking, the employment of mathematically ill-defined concepts within the context of what I will call an 'inferentially restrictive' methodology. (This is intended to include the employment of mathematically self-contradictory concepts.) I will show how the physicist's use of delta functions and path integrals may be interpreted as violations of rigor of this sort that nevertheless contribute to physically persuasive arguments. This will help to demonstrate that the liberal view is indeed implicit in the methodology of physics as it has historically been practiced.⁸ Merely showing that the liberal view is implicit in the physicists' practices, however, still leaves open the *normative* question of whether the

⁶On pp 93-94 of [9], for instance, Feynman distinguishes mathematical rigor from physical rigor, asserting that they do not necessarily coincide. This attitude towards mathematical rigor can also be found throughout his famous undergraduate texts [10].

⁷For a discussion of Bohr's aversion to mathematics, see pp259-262 of Beller [1].

⁸Of course, this is not to say that all violations of mathematical rigor found in physics are of this sort, or that any violation not of this sort cannot be defended. To the contrary, I am quite convinced that there are radically different sorts of transgressions of mathematical rigor whose use in physics can be defended – but not necessarily by the arguments that will be given in this paper. In this paper I have simply tried to isolate one important class of violation, and defend it in a way that usefully illuminates some interesting problems about the relation between mathematics and the physical world.

physicist is *entitled* to behave in this way. In the third section, I will argue that the practice of physics is not misguided in this way, and that the physicist is indeed entitled to pursue a methodology in which he permits himself to be moved by arguments violating the usual standards of mathematical rigor in the specific way described. In the fourth section, I will then dispel some general worries one might have about the employment of mathematically unrigorous methods in physics. In doing so, I aim to show that the demands placed on the physicist by the conservative and moderate views defined above are not reasonable.

2 Rejecting Inferential Permissiveness.

In this section, I wish to identify and discuss a specific class of unrigorous argument that one generally finds tolerated in the physics literature. The type of reasoning I am interested in involves concepts with ambiguous, or even self-contradictory properties. Specifically, I am interested in cases in which such concepts are employed by physicists with the proviso that certain otherwise mathematically acceptable forms of argument involving these concepts be prohibited. Prohibitions of this sort act as a way to control any damage that might arise from the ill-definedness or inconsistency of the concept in question. What makes this phenomenon interesting is that a restriction of this sort *cannot* be justified from the mathematician's point of view. The mathematician cannot refuse to apply an otherwise acceptable form of argument to a proposition merely because of the proposition's content. One of my goals will be to understand how what is unthinkable for the mathematician in this way can be accepted so easily by the physicist.

I will focus on two examples of this phenomenon - the physicists' use of the delta function, and the use of path integral measures in quantum mechanics. Before beginning a more detailed discussion of these examples, however, it will be helpful to characterize the phenomenon in which I am interested in more general terms.

When the mathematician tries to prove a theorem or solve a mathematical problem, he is free to invoke any concept, technique or theorem from any area of mathematics that he wishes. No *legal* inference from true premises is barred – although to say an inference is not barred is obviously not to say that it bears fruit. It is precisely because of this 'permissive' nature of inference in mathematics that one comes across surprises such as the set-theoretic proof of the number-theoretic claim that transcendental numbers exist. By describing mathematics as *inferentially permissive*, I shall mean that there is no restriction on the concepts or legal rules of inference that one can invoke in attempting to prove a theorem or solve a puzzle in mathematics. In a mathematical proof, any legal inference is

permitted and any concept can be invoked, regardless of the subject matter introduced.

Because mathematics is inferentially permissive, when we use the language of mathematics to describe problems in physics, we open the door to the possibility that unexpected areas of mathematics might turn out to be relevant to physics. Excellent examples of this are the application of group theory to modern particle physics, and the use of the theory of analytic functions in quantum field theory. The possibility of such surprises is surely one of the reasons that the physicist goes to such great lengths to characterize in *mathematical* terms the phenomena in which he is interested.

Yet in spite of the potential rewards of exploiting the inferentially permissive character of mathematics in physics, a closer look at the methodology of physics shows, I claim, that the physicist is *not* willing to consider absolutely *any* sort of mathematical technique or rule of inference when solving a problem. His willingness to exploit the inferentially permissive character of mathematics, I shall argue, only goes so far. I shall summarize this by claiming that the methodology of physics is *inferentially restrictive*. Much of this section will be devoted to showing that this is indeed a true claim about physics as we find it practiced.

Having introduced (but not defended) the distinction between the inferentially permissive methodology of mathematics and the inferentially restrictive methodology of physics, let us now return to question of rigor in physics.

One might think that if, as I have asserted, the methodology of physics is inferentially restrictive, then physics must be all the *more* rigorous for it – after all, when we think of nonrigorous reasoning, we generally think of reasoning which goes *beyond* the argument forms generally accepted by the mathematician. Surely being more restrictive than the mathematician could only improve things.

But not all unrigorous reasoning need involve such things as explicit violations of the rules of the predicate calculus. One can sometimes engage in reasoning with a self-contradictory concept or inconsistent set of premises by deliberately abstaining from rules or patterns of inference which would obviously lead from the concept or premises in question to a contradiction. Insofar as such reasoning involves concepts or premises which *could* lead to contradictions in fairly obvious ways were certain rules of inference to be exploited, the mathematician would not describe such reasoning as rigorous, regardless of whether the reasoner explicitly abstains from such rules of inference or not. In such cases, one reasons unrigorously not so much by violating the laws of the predicate calculus, but rather by refusing to take such laws seriously enough – specifically, one reasons unrigorously insofar as one employs concepts and premises in a way which overtly flies in the face of such things as no-go or impossibility theorems. The earliest attempts at exploiting infinitesimals, for example, were criticized as

unrigorous precisely on these grounds, and generally not on the grounds that facts such as the law of non-contradiction had been violated.

The sorts of unrigorous arguments in physics I wish to consider are precisely of this sort. Such violations of mathematical rigor are possible only insofar as the methodology of physics is inferentially restrictive. Thus, if it is true that the type of nonrigorous reasoning in question is implicit in the practice of physics, it follows that the practicing physicist adopts an inferentially restrictive methodology.⁹

Note, however, that I need not claim that *all* instances of nonrigorous argument in physics are of this form. In order to justify the liberal viewpoint, it suffices to show that there are *some* mathematically unpersuasive arguments which ought to be regarded as persuasive pieces of physics. Different types of nonrigorous arguments in physics may well require different justifications; and there are perhaps arguments in physics whose nonrigorous nature cannot be justified at all. But so long as the rejection of inferential permissiveness is regarded as a legitimate methodological tool in physics – even in the case of mature theories – there will be a type of nonrigorous argument in physics whose nonrigorous nature *can* be justified. This is enough to prove the liberal view correct.

Let us now turn to my primary examples of the phenomenon in question – the physicists’ use of delta functions, and the use of path integral measures in quantum mechanics.

In many expositions of quantum mechanics, the delta function $\delta(x)$ is used to define wavefunctions of localized states in the position representation. It satisfies the defining relations:

$$\int_{-\infty}^{+\infty} \delta(x) dx = 1 \quad \text{and} \quad \int_{-\infty}^{+\infty} f(x) \delta(x) dx = f(0)$$

for any L^2 function f .¹⁰ It is easily shown, however, that no such function $\delta(x)$ exists. For let f and g be L^2 functions such that $f(x) = g(x)$ almost everywhere, but $f(0) \neq g(0)$. Then, since $f(x)\delta(x) = g(x)\delta(x)$ almost everywhere,

$$\int_{-\infty}^{+\infty} f(x) \delta(x) dx = \int_{-\infty}^{+\infty} g(x) \delta(x) dx.$$

But the left hand side of this equation is just $f(0)$, and the right hand side $g(0)$, contradicting $f(0) \neq g(0)$.

Delta functions are involved in many calculations in quantum mechanics. Given that the delta functions are, at best, a fiction, one might then wonder why we should trust the calculations of quantum mechanics involving them. Happily, there is a straightforward

⁹I also think that there are instances of inferential restrictiveness in physics that have nothing to do with questions of rigor, but those will not be the subject of the present paper.

¹⁰A complex-valued, measurable function f is L^2 just in case $\int_{-\infty}^{+\infty} |f(x)|^2 dx < \infty$.

answer – so long as one keeps all delta functions inside integral signs, one cannot generate a contradiction.¹¹ Fortunately, this solution applies to quantum mechanics. In calculating, for instance, the expectation value of an observable for a state localized at a given point, all relevant delta functions *do* occur within integral signs, and the result obtained is not mere nonsense.¹² As long as the physicist never asks questions about delta functions occurring outside integral signs – for example, as long as he never asks for the value of the delta function at a given point – his calculations will not be affected by mathematical contradictions such as the one identified in the previous paragraph.

It might be suggested, however, that in contrast to what I have said, the *real* reason that delta functions do not lead us into disaster is *not* that physicists knowingly refrain from interrogating delta functions outside integral signs, but rather that delta functions may be treated as a more abstract class of mathematical entity known as *distributions*.¹³ Once we abandon the idea that the delta function is actually a function and embrace the idea that it is a distribution, the threat of inconsistency vanishes. This, it might be suggested, is what saves the calculations of quantum mechanics from contradiction.

If one thought, however, that this was the real reason that delta functions do not lead physicists into disaster, one would have to be surprised – and perhaps even a little perplexed – by the fact that most quantum mechanics texts use delta functions extravagantly, without explaining in any satisfactory level of detail the fact that they are to be properly understood as distributions rather than functions. If it is this critical fact about delta functions that saves almost every quantum-mechanical calculation from meaninglessness, one would expect the correct interpretation of delta functions to be a matter of great urgency in the physics texts. But no such sense of urgency is to be found. What one finds, instead, is a curious sense of indifference. Because of this, I do not think that the existence of distribution theory on its own can fully account for the physicist’s nonchalant attitude towards delta functions.

In fact, a closer look at the physics literature shows that the rigorization of delta functions

¹¹I will not provide the details of the argument for this here, but note that the arguments just given for the non-existence of the delta function revolve around considering the properties of the delta function outside the context of an integral.

¹²For instance, the expectation value for the position observable \hat{X} , when the system is in the state given by the wavefunction $\delta(x - c)$ (in the position representation), is just

$$\langle \psi | \hat{X} | \psi \rangle = \int_{-\infty}^{+\infty} \delta^*(x - c) x \delta(x - c) dx = \int_{-\infty}^{+\infty} \delta^*(x - c) c \delta(x - c) dx = c \int_{-\infty}^{+\infty} |\delta(x - c)|^2 dx = c.$$

Nothing in this calculation requires that we know how the delta function behaves outside the context of an integral.

¹³To be fully rigorous about the use of delta functions in quantum mechanics, one would also have to invoke such things as spectral theory, as demonstrated by von Neumann in [25]. The full details of this are not important for what follows, however.

by Schwartz in his 1945 paper [21]¹⁴ was an event of little significance for physicists, who generally felt long before then that the situation with delta functions was already well under control.¹⁵ For instance, Dirac’s textbook [6] – one of the classic texts of early quantum mechanics, first published in 1930 – showed very little anxiety over the use of the delta function, in spite of the fact that Dirac took the unsoundness of the delta function as a mathematical entity to be perfectly obvious to everyone. In this work, Dirac’s only real comment indicating that there was something unusual about the delta function occurred in §15, where he wrote:

*‘... $\delta(x)$ is not a function of x according to the usual mathematical definition of a function, which requires a function to have a definite value for each point in its domain, but is something more general, which we may call an ‘improper function’ Thus $\delta(x)$ is not a quantity which can generally be used in mathematical analysis like an ordinary function, **but its use must be confined to certain types of expression for which it is obvious that no inconsistency can arise.**’* (Emphasis mine.)

These few remarks suggest that once one commits to avoiding the delta function in manifestly troublesome contexts – a commitment which Dirac thinks can be honored in quantum mechanics – no questions about the the delta function worthy of the physicist’s attention remain. For Dirac, then, it is going inferentially restrictive that resolves any serious problems surrounding the use of the delta function. Furthermore, this view seems to be quite typical of the physics literature of the time. In light of this, it should be no surprise that the work of Schwartz had little effect on physicists, almost all of whom – much like Dirac – felt that there was really no great problem about the delta function, once the commitment to an inferentially restrictive methodology was made.

In no way did physicists change their tune after Schwartz’s work of 1945. Even after the advent of distribution theory, most physicists were still happy to stick with the thought that it is judicious use of the delta function, rather than the existence of distribution theory, that saves quantum mechanics from disaster. On p67 of Dick and Wittke’s 1960 textbook [8], we are told, for instance:

‘The Dirac delta function is meaningful only under integral signs, where the limiting technique can be used.’

¹⁴For a little more detail on the history of the rigorization of delta functions, see pp 313-314 of Jammer’s [13].

¹⁵It should be noted that the mathematical physicist Von Neumann was an exception to this. Von Neumann felt the introduction of delta functions introduced ‘great mathematical difficulties’ ([25], page 31) that he took very seriously. This, however, was a minority view.

Comments like this continue to pervade physics texts even to this day, in spite of the fact that a rigorous treatment of the delta function is now available. One must be struck by how little the lessons of distribution theory have impacted the physics community. Even Messiah's textbook – which actually contains a brief appendix on distribution theory – did nothing to challenge the view that the unrigorous, inferentially restrictive approach to delta functions was perfectly adequate. In Chapter V §8, for instance, Messiah tells us:

‘... In fact, no such function of p' and p'' possesses the desired property. Nevertheless, if one does not bother too much about mathematical rigor, one can, following Dirac, make use of the “singular function” $\delta(x)$ defined by the property ...’

where one of the usual definitions follows. Moreover, Messiah's appendix on distribution theory – one of the only in the quantum mechanics literature – is so cursory that one can hardly imagine it being useful to anyone seriously interested in the project of a rigorous reconstruction of quantum mechanics free from the delta function. So, in spite of Messiah's explicit acknowledgement of distribution theory, one cannot extract from his book the lesson that distribution theory is something that must be studied by any quantum theorist with a conscience who wants to sleep soundly at night.

In fact, the only thing that seems to separate the more careful from the less careful quantum mechanics texts is *not* the extent to which they discuss the details of the theory of distributions, but rather the extent to which they at least make some effort to inculcate the right set of *commitments* about the use of delta functions. No fundamental change occurs when one looks at papers published in the physics journals, instead of texts – the attitude to the delta function there remains the same. So even though after 1945 physicists were *able* to take an inferentially permissive approach to the delta function by invoking distribution theory, they nevertheless generally remained quite content to stick with their naive, unrigorous conception of the delta function, and pay the price of inferential restrictiveness. This attitude, which had been established well before 1945, was essentially unaffected by the discoveries of Schwartz.

All of the above suggests that in the physicist's view, the rigorization of the delta function was simply irrelevant to the question of the soundness of the usual calculative techniques of quantum mechanics. (Of course, the claim might be made that physicists were *wrong* in caring so little about the foundational issues surrounding the delta function. We shall turn to this issue later.) The use of the delta function in quantum mechanics shows that physicists are sometimes perfectly happy to knowingly exploit self-contradictory concepts –

even when rigorous alternatives are known to be available. The physicist knows, of course, that this compromises his ability to adopt an inferentially permissive methodology. But the lack of interest in questions about the rigorization of the delta function seems to indicate that the loss of inferential permissiveness – at least in this case – is by no means devastating for the physicist.

Our next example of unrigorous reasoning in physics is the use of the path integral, which was invented and used by Feynman in order to explain a new way of thinking about the basic concepts of quantum mechanics. If a and b are space-time points, and $K(a, b)$ is the amplitude for a particle to go from b to a in the quantum system whose classical analogue has Lagrangian \mathcal{L} , then in the path integral formalism the identity

$$K(b, a) = \int_a^b e^{\frac{i}{\hbar} \int_a^{t_b} \mathcal{L} dt} \mathcal{D}x(t)$$

holds, where $\mathcal{D}x(t)$ is the ‘path integral measure’ over the set of possible paths from a to b through classical phase space. The path integral method gives an entirely different way of doing quantum-mechanical calculations from the usual operatorial methods. It may be given an intuitive motivation that the operatorial approach to quantum mechanics lacks. It also has the added feature of making certain numerical calculations in quantum mechanics more tractable.

It is unclear, however, how to construct the underlying measure $\mathcal{D}x(t)$ relative to which the path integral is to be defined. In the case of imaginary time (i.e., Euclidean space), the path integral measure can be defined with the use of the Wiener measure, a measure used originally in the study of Brownian motion. (See Glimm and Jaffe’s [11] for details of this.) Of course, in field theory, there are many situations in which useful calculations *do* involve imaginary time, but for a completely general treatment of path integrals, the Wiener measure is insufficient. In complicated situations, such as when one uses the path integral to describe interacting fields in Minkowski space (rather than Euclidean space), there is little guarantee that there really is a well-defined underlying measure that can be used in the path integral. In fact, there are even theorems showing that in certain cases one *cannot* define a path-integral measure satisfying the requirements of the physicist.¹⁶

Here, we cannot simply gesture to some mathematical theory the way we can gesture to the theory of distributions when questioned about delta functions. The question of how the path integral is to be understood in full generality remains open. Given this, one might expect to see the physicist expending great energy trying to clarify the precise mathematical

¹⁶Cameron [5] shows that there are cases in which no measure exists that can be used to define the Feynman integral. For a discussion of this and other difficulties involved with the path integral, see Chapter 6 of Rivers [20].

meaning of the path integral. Curiously, we again find that this is not the case. The typical physicist shows little interest in the rigorous work mathematicians have done on the path integral. This situation bears a striking resemblance to the pre-1945 attitude towards the delta function – physicists simply see no point in losing sleep over a problem that they largely take to be well under control.

But what makes the physicists think that the situation with path integrals *is* largely well under control? The answer is that as long as the uses to which the path integral are put are of a particular sort, problems of inconsistency can be avoided. The pre-1945 use of the delta function was generally regarded as unproblematic so long as the delta functions were only used in a particular context – i.e., in expressions to be used as integrands. Likewise, there are ways of limiting one’s use of the path integral in such a way that the threat of inconsistency disappears.

One can, for instance, think of the path integral formalism as merely an efficient algorithm for the generation of perturbation theory in either quantum mechanics or field theory. If we treat the path integral formalism in this way, no threat of inconsistency arises (so long as we take for granted the consistency of perturbation theory itself). Furthermore, this treatment of the path integral formalism does not demand that we take the path integral measure $\mathcal{D}x(t)$ seriously, just as employing the delta function within an integrand does not presuppose that we take the delta function seriously as a function.

In addition to this, provided that the path integral satisfies the usual properties of an integral, along with the ‘convolution’ property

$$K(b, a) = \int_{x_c} K(b, c) K(c, a) dx_c,$$

one can actually re-derive the Schrödinger equation.¹⁷ These properties of the path integral on their own are certainly consistent, and – because they may be used to derive the Schrödinger equation – are sufficient for doing the non-relativistic quantum mechanics of particles. This provides us with yet another way of demarcating the argument forms involving path integrals that may be used without fear of inconsistency, in spite of the meaninglessness of the path integral measure $\mathcal{D}x(t)$.

There is, however, an important point of difference between the delta function and path integral. Unlike the pre-1945 use of the delta function, it is unclear *exactly* how literally one can take the idea of a path integral without getting into trouble. We know that the path integral may be used unproblematically to generate perturbation theory, and we know that the path integral cannot be taken completely literally as an integral supported by a

¹⁷See §4.1 of Feynman and Hibbs [9].

well-defined underlying measure. There are, however, many ‘degrees of seriousness’ between these extremes with which one may want to take the idea of the path integral, and for which it is generally *not* clear whether one is on safe ground or not. If the physicist wants to go beyond the safest ways of treating the path integral, he must make conjectures about consistency which he may well be forced to rescind at some later point. These conjectures are necessary because the conceptual and mathematical problems behind the path integral are just not as transparent as those with the delta function were prior to 1945. Thus, in the case of the path integral, the precise bounds of the physicist’s inferential restrictiveness are not entirely clear. Yet in spite of this, the basic phenomenon in which I am interested remains unchanged – the physicist adopts a *mathematically incoherent* concept, and in order to avoid inconsistency, accepts certain inferential restrictions. The claim that the inferential restrictions in question successfully manage to avoid inconsistency may be a matter of conjecture rather than certainty, but the methodological fact that one permits mathematically incoherent concepts into one’s discipline by going inferentially restrictive remains unchanged.

In summary, my central claim about the use of the path integral in physics is much the same as my claim about the use of the delta function in physics – it is a type of unrigorous argumentation which does not bother the physicist, insofar as he is happy to adopt an inferentially restrictive methodology that protects him from unpleasant consequences.¹⁸ The willingness to go inferentially restrictive in turn explains the indifference one finds amongst practicing physicists to the progress the mathematician has made or continues to make in rigorizing the concept of the delta function or path integral.¹⁹ Insofar as physicists are happy to be moved by arguments which are neither rigorous nor obviously rigorizable – even in a highly mature discipline such as non-relativistic quantum mechanics – it is clear that the liberal view outlined in the previous section does indeed permeate the practice of physics.²⁰

In this section, I have been content to make descriptive claims about the practice of

¹⁸In [18], Norton has suggested that something very much like adopting an inferentially restrictive methodology may be useful for sidestepping certain inconsistencies in Newtonian Gravitation Theory. For a careful discussion of the inconsistency in question, see also Malament’s reply [17] to an earlier version of Norton’s paper.

¹⁹Of course, in the sense that the mathematics literature might help to validate or refute the physicist’s conjectures about the consistency of certain argument forms involving path integrals, the physicist might have something to gain from the mathematician. This, however, is generally not the direction of mathematical research into the path integral.

²⁰It is important to note that the inferentially restrictive methodology of physics does not make it impossible for the blind pursuit of mathematics to lead to physically significant discoveries. A good example of progress in physics being made in this way would be the discovery of anti-particles associated with negative energy solutions of the Dirac equation. Other examples of discoveries in physics arising from purely formal or mathematical considerations are discussed by Steiner in [24]. Nothing about the adoption of an inferentially restrictive methodology prevents novel applications of purely mathematical ideas in physics; even though their creators will perhaps be required to defend such ideas with greater vigor than they would were the methodology of physics inferentially permissive.

physics. A natural response to all this, however, is that in spite of actual practice, physicists should *not* be persuaded by unrigorous arguments – especially in mature theories – and likewise should *not* be so nonchalant about adopting an inferentially restrictive methodology. In order to defend the liberal view against these attacks, we must now turn our attention to normative issues such as these.

3 The Agreement Problem.

In the previous section I argued that the practice of physics is inferentially restrictive, and that this is how a certain type of non-rigorous argument finds its way into physics. This, however, is just a descriptive claim that leaves unanswered the question of whether physicists are *entitled* to employ an inferentially restrictive methodology. In this section, I will address this question, answering it affirmatively.

Let us consider again the role of mathematical reasoning in physics. The possibility of using mathematical argumentation to derive conclusions in physics relies on representing physical entities as if they were mathematical entities. For instance, we treat the trajectories of particles as curves whose second time derivatives are everywhere defined, we think of bodies as having symmetries that are given by some group, and so on. This process of treating physical entities as if they were mathematical entities I shall call *mathematical interpretation*. Without mathematical interpretation, mathematics could not make the deep contact with physics that it does today.

The process of mathematical interpretation involves philosophical difficulties that might not be immediately apparent. Consider, for instance, the treatment of the trajectory of a classical particle as if it were a curve whose second time derivative is everywhere defined. This interpretation commits us to the claim that, at each time, some number occupies the 200th decimal place of the x co-ordinate of the particle. A physicist would be extremely cautious, however, about using this fact in any calculation of physical significance. Specifically, many physicists would quietly wonder whether the 200th decimal digit of a spatial co-ordinate really has any physical meaning at all, even though the formalism of real numbers – which is utterly indispensable for classical mechanics – entails its existence. When the choice is made to mathematically interpret the idea of ‘location’ with the real numbers, the physicist does not intend to exploit the fact that such numbers have a 200th digit in their decimal expansion. Nevertheless, the physicist is, at the very least, ‘mathematically committed’ to figures in the 200th decimal place, even if he remains agnostic about their physical meaning. Given this, to exploit such features of the real numbers in physics arguments would be to

take the act of mathematical interpretation more literally than it was ever intended to be taken. This, in turn, invites mistrust of any results obtained.

Consider another example. For any classical trajectory, we can define the cardinality κ of the set of points at which its tenth derivative is defined. Now it becomes possible – at least, in principle – to use the cardinal κ in an argument whose conclusion may have some particular physical consequence. But again, no physicist would accept any such conclusion without a great deal of caution. When the choice is made to mathematically interpret the trajectory of a classical particle as a twice differentiable curve, the physicist’s intention is to allow himself to exploit the mathematics of curves most directly related to their being twice differentiable. To exploit unrelated features of such curves is again to take the act of mathematical interpretation more literally than it was intended to be taken, and invites similar mistrust.

Mathematical interpretation often threatens to bring ‘excess baggage’ with it in precisely this way. Some of the epistemological puzzles related to such excess baggage, along with some helpful examples of this phenomenon, have been discussed by Mark Wilson [26, 27]. In practice, the way the physicist ensures that excess baggage does not interfere with his results is by restricting the sorts of mathematical arguments invoked when solving a physics problem. By adopting an inferentially restrictive methodology, the physicist ensures that what he judges to be mathematical irrelevancies do not affect his calculations.²¹ He shuns problematic argument forms such as those given in the previous paragraphs, and instead sticks to a more limited set of argument forms, about whose physical meaning he feels most confident. In this way, the physicist justifies his use of an inferentially restrictive methodology.

Still, it might be said that this does not address the *normative* question of whether the physicist should at least *aspire* to pursue an inferentially permissive methodology. It might be claimed, for instance, that at least one of the goals of physics should be to discover a mathematical interpretation of the concepts of physics for which there is no threat of having to reject any mathematically valid argument on the grounds that it involves spurious mathematical considerations. Let us then define the *agreement problem* as the problem of deciding if there is a mathematical interpretation of physical reality in which no mathematically persuasive argument can take us from true premises to a false conclusion. If we decide that such a mathematical interpretation of physical reality exists, the explicit construction

²¹Of course, in some cases it can be difficult to tell which sorts of mathematical notions are physically relevant and which are not. The careful physicist may at times have to change his assessment of what sorts of mathematical judgements may be dismissed as physically irrelevant. This complication, however, does not affect the point in question.

of this interpretation will also be regarded as part of the solution of the agreement problem.

Note that if we were able to construct a mathematical interpretation of physical reality settling the agreement problem in the affirmative, we would then have the liberty of treating any physics problem as if it were a purely mathematical problem, with no need to worry about the physical meaning or significance of any of the mathematical entities involved, and with no threat of taking any act of mathematical interpretation more literally than it was intended to be taken. In such a case, we would be free to adopt an inferentially permissive methodology. Our question then becomes why we should not view the solution of the agreement problem – and in particular, the construction of an interpretation settling it in the affirmative – as part of the job of the physicist. If it were decided that this was indeed part of the job of the physicist, then the liberal view could not be defended by merely pointing to the existence of inferentially restrictive methods in physics. The reason for this is that if we take it to be one of the ultimate goals of physics to solve the agreement problem in the affirmative, then the existence of inferentially restrictive methods can be straightforwardly dismissed as nothing more than the sign of an immature physics – immature insofar as it has not yet affirmatively solved the agreement problem!

I claim, however, that it is not appropriate to think of the construction of a mathematical interpretation of physical reality settling the agreement problem in the affirmative as a *necessary* part of physics. To begin to see why, it is first helpful to be reminded that even in the simple case of the classical mechanics of point particles, the difficulties involved in solving the agreement problem are substantial.²² If our prospects for solving the agreement problem remain so bleak in a discipline that we take ourselves to understand so well *physically*, we must begin to wonder whether or not it is really fair to view the settling of the agreement problem in the affirmative as a necessary part of physics, without which physics would be left incomplete.

In fact, I think that physicists are generally agnostic about whether an affirmative solution to the agreement problem exists at all, and regard the question of whether such a solution exists as something of little relevance to the needs and goals of physics. But the claim that I must presently defend is not the weaker claim that the practicing physicist regards the settling of the agreement problem as irrelevant to his work, but rather the stronger claim that he is *entitled* to take this view.

So *is* the physicist entitled to remain agnostic about the existence of an affirmative solu-

²²Wilson [26, 27] outlines just some of the conceptual difficulties involved. For instance, one must view the mathematical representation of the boundary of a liquid as having sufficient structure that it be the sort of thing that can be propagated forward in time by a differential equation, but not so much structure that local changes in the curve cause instantaneous global changes elsewhere on the boundary, due to analyticity or Bezout's theorem considerations.

tion to the agreement problem? After all, it could be claimed – contrary to what I have been suggesting – that it is simply *obvious* that the agreement problem has an affirmative solution, and therefore also obvious that the physicist is *not* entitled to remain agnostic about the existence of an affirmative solution to the agreement problem. It is beyond dispute, after all, that the physical world has some sort of mathematical structure demanding elucidation. It is also beyond dispute that this elucidation must form an integral part of the physicist's broader project of articulating the laws of nature. If we take this to mean that the agreement problem has an affirmative solution that the physicist must strive to discover, then agnosticism about the agreement problem cannot be a legitimate option for the physicist. While it might be legitimate for the physicist to wonder whether a formalism depending on the real numbers could form part of the affirmative solution to the agreement problem, it would not be legitimate for him to wonder whether the agreement problem has a solution *at all*. According to this view, that the agreement problem is solvable in the affirmative is to be taken as one of the very premises, or articles of faith, of physics itself.

This way of thinking about physics, however, takes too much for granted. While it does seem reasonable to conjecture that the world has some sort of deep mathematical structure, it is far from obvious that the world has a mathematical structure which is compatible with an inferentially permissive methodology. The *only* thing the physicist takes as obvious is that there are regularities and rhythms underlying the world which might well be elucidated with the help of the language and methods of mathematics. It is difficult to see how the physicist's attempts to understand these regularities and rhythms presuppose any stance on the agreement problem. The successes of classical mechanics, for instance, show that it is possible to understand and articulate the rhythms of nature without taking any real stance on the agreement problem.

In brief, I do not see how the very idea of physics presupposes that the agreement problem can be solved in the affirmative. All the usual questions of physics can be motivated and pursued *without* taking a stance on the agreement problem. Furthermore, there are great difficulties associated with any attempt at constructing a mathematical interpretation of physical reality settling the agreement problem in the affirmative. These facts, I claim, entitle the physicist to remain agnostic about the agreement problem. Indeed, they entail that he positively *ought* to remain agnostic about the agreement problem, insofar as to do otherwise would be to take an unnecessarily bold stance on an issue which, at present, the physicist has no need or reason to take a stance on, and which, if the wrong side were taken, could severely impede the development of physics. In such a situation, agnosticism is not just an intellectual possibility - it is an intellectual duty.

Let us summarize. The physicist is entitled to remain agnostic about the agreement problem. Consequently, when he mathematically interprets his physical concepts, he must realize that applications of arbitrary mathematical techniques need not take true premises to true conclusions. He will acknowledge the possibility that his interpretation only remains valid as long as certain types of mathematical techniques are used, and others avoided. Such a physicist will therefore be open to the possibility that inferentially restrictive methods are needed in physics. But the employment of inferentially restrictive methods is precisely what allows the physicist to introduce mathematically incoherent concepts into his discipline. Therefore, the physicist will be forced to acknowledge the possibility that mathematically incoherent concepts may be necessary in physics. It follows that the liberal view about the role of mathematical rigor in physics must be correct.²³

This concludes the argument of this section, and with it, my defense of the liberal view. I wish now to address two worries about the details of the material presented in this section. In the next and final section of this paper I will then discuss some of the more general concerns that one might have about the admission of mathematically unrigorous methods into physics.

The first worry I wish to address is that insofar as the physicist pursues an inferentially restrictive methodology, he does *not* remain agnostic about the agreement problem, but rather assumes that it *cannot* be solved affirmatively. To see why, note that in cases such as those presented in the previous section, we have mathematically rigorous arguments which we *must* reject as physically misleading – such as our deduction of a contradiction from the definition of the delta function – lest all hell break loose. But if in such cases we are forced to declare that a mathematically rigorous argument is physically irrelevant, then it would seem that we are no longer able to hold out any hope that the agreement problem might be solved affirmatively – for recall that if the agreement problem can be solved affirmatively, then *no* mathematically valid argument can ever be dismissed as physically irrelevant. Hence we can no longer be said to be agnostic about the agreement problem. But this criticism is not valid. If one were to claim with certainty that *no* inferentially permissive methodology could ever produce the same conclusions as the ones we get with the use of delta functions or path integrals in quantum mechanics, one could certainly not profess to be agnostic about the agreement problem. But when delta functions and path integrals are used, no physicist ever makes any such claim – indeed, in the case of delta functions, such a claim would be

²³Of course, none of this is intended to suggest that mathematicians, philosophers, and even physicists of a certain persuasion should not try to study the agreement problem, and consider the extent to which an inferentially permissive methodology might be possible in physics. The point is rather that the completion of such a study – if it is possible at all – should not be thought of as a necessary part of physics itself.

false, as the proponents of distribution theory are keen to remind us. In these cases, it is insofar as the physicist shows little concern over the question of whether his arguments are *rigorizable* – rather than rigorous – that his agnosticism about the agreement problem rears its head.

The second worry is that although the physicist who seeks to take no sides on the agreement problem will have to remain open to the possibility of adopting an inferentially restrictive methodology, this does not seem to be enough to justify the rejection of any *specific* form of mathematical argument. It is one thing to know that an inferentially restrictive methodology may be necessary in physics, but it is quite another to be justified in rejecting any *particular* form of mathematical argument. Until we have some way of justifying such rejections, the knowledge that such rejections may be necessary is of little use.

But this worry is also unfounded. Such justifications are not too hard to come by. I have already referred to one type of justification that might be given, and there are other types of justification still. Recall that when an act of mathematical interpretation occurs, the physicist generally has the application of some type of mathematical technique in mind. For instance, in classical mechanics, we model trajectories as twice differentiable curves precisely so we can exploit the integral and differential calculus. Given this, it becomes possible for the physicist to notice that a mathematical trick or technique applied in the context of a physics argument goes way beyond the sorts of techniques originally intended when the relevant mathematical concepts were introduced. Insofar as we acknowledge the possibility that an inferentially restrictive methodology might be necessary for physics, we are surely justified in such cases to at least *conjecture* that the mathematical trick or technique in question may be misleading, and that some sort of inferential restriction may be warranted. Of course, such a conjecture is not obligatory, and even when it is made, it can later be retracted or modified. But this does not make such a conjecture unjustified. Indeed, if we think it is possible that the agreement problem cannot be solved affirmatively – as the person who seeks to take no stand on the agreement problem must – then we would be remiss if we never made such conjectures, insofar as it is possible that physics ultimately *demand*s an inferentially restrictive methodology.

In the examples of the previous section, we find precisely such conjectures being made. The delta function is introduced into quantum mechanics in order to represent eigenvectors of an operator with a continuous spectrum, and in no way is it intended to be the sort of thing whose properties outside of an integral should be useful. Likewise, the path integral measure is introduced in order to help calculate quantum-mechanical transition amplitudes, and in no way is it intended to be useful in any other context. In both cases, it is conjectured that the

relevant troublesome argument forms are too deviant to be taken seriously, and consequently should be ignored. In the former case, the conjecture has been proven correct, insofar as the calculus of distributions allows us to formulate the concepts of quantum mechanics in a way that eliminates the troublesome argument form. In the case of the path integral, the conjecture that it is safe to ignore the relevant troublesome argument forms is also reasonable, although many questions in this area still remain open.

4 Independent Objections to the Liberal View.

In this final section, I wish to consider three arguments that might be given against allowing physicists to be moved by mathematically unpersuasive arguments, and show how they may be defused. According to the first objection, employing nonrigorous mathematics compromises our ability to *understand* why the calculations of physics work. The second objection I will consider is that nonrigorous mathematics compromises our ability to view our conclusions as *justified*. The final objection I will consider is that once we allow any non-rigorous mathematical methods into physics, we deprive ourselves of the ability to criticize blatantly unacceptable abuses of mathematics in physics.

The first worry is that as long as we permit the use of mathematically unpersuasive arguments in physics, we can never *understand* why the calculations of physics can be so effective, when in fact they are. To motivate this concern, recall that in [14], Kitcher argues that one of the things that makes rigor necessary in mathematics is that it helps us to understand why our mathematical calculations work. Here, Kitcher is thinking of the calculations that lie behind *mathematical* problem solving, such as finding roots of equations, calculating maxima and minima, etc., rather than problem solving in physics. But there is an obvious analogue of Kitcher's position in the world of physics problem solving, according to which rigor is necessary in physics precisely because it helps us to understand why our *physics* calculations work. If rigor is required in physics for this reason, then the liberal view cannot be acceptable.

To address this concern, let us ask what is meant when it is said that we do (or do not) understand why a calculation works. One can imagine a calculation involving strange steps, in which it is simply not clear how to view the argument as exhibiting the way in which the laws of nature, as they are ordinarily understood, lead to the result or phenomenon in question. In such cases we confess that we do not understand why the calculation works. By contrast, there are calculations in which each step follows inevitably from the one before, and the calculation as a whole can be viewed as a natural expression of the way in which

simple, well-understood, processes come together to produce the physical result in question. In cases such as these, we claim to understand why the calculation works.

To what extent, then, is mathematical rigor needed to produce this feeling of understanding? There is no denying that sometimes attention to rigor can help to produce understanding. There is also no denying that, when abused, nonrigorous mathematical methods can hinder understanding. But these are nothing more than rules of thumb. Rigorous methods can sometimes be entirely unilluminating, and unrigorous mathematical calculations can sometimes bring with them a deep understanding of how the underlying processes work together to produce the phenomenon under investigation. In the case of nonrigorous reasoning, part of this understanding will involve appreciating that certain sorts of mathematical interrogation should not be regarded as physically significant, and should consequently be ignored. It is difficult to see how this hinders understanding. The case of path integral methods in quantum mechanics provides a good example of this.

My opponent could assert that whenever inferences which would lead to unpalatable conclusions are dismissed as physically irrelevant, understanding is necessarily compromised. In the previous section, however, I have argued that one can actually justify employing an inferentially restrictive methodology in virtue of not wanting to take a position on the agreement problem. My opponent would therefore have to claim that this sort of agnosticism is incompatible with the desire for understanding. But this claim seems like an obvious exaggeration. I do not think that the connection between rigor and understanding in physics is as deep as this first worry suggests.

The second worry about the use of nonrigorous methods in physics is closely related. According to this worry, we can never view the result of a calculation that employs non-rigorous methods as *justified*.

Note, however, that in an inferentially restrictive methodology, each step of every argument is still logically valid. What distinguishes inferentially restrictive methodologies is the commitment to *avoid* certain otherwise legal inferences – but that does not mean that the inferences one actually performs are illegal. There is no question, then, that the conclusion of a calculation or argument is entailed by the premises in the sort of nonrigorous argument I have been focusing on. In this sense, there is no problem of justification.

This could be granted, but my opponent might object that we are never justified in ignoring an otherwise valid logical inference. So even though a calculation might lead to a true conclusion, the fact that if certain barred inferences were to be permitted, a contrary result would be obtained, is enough to make it the case that the calculation in question does not really justify the answer. One can, after all, prove any conclusion one wants by starting

with an inconsistent set of premises and cleverly restricting inferences.

Of course, those who work in paraconsistent logic would argue that one *can* justify a conclusion from an inconsistent set of premises. In many obvious ways, reasoning in an inferentially restrictive manner is like paraconsistent reasoning.²⁴ The existence of approaches to logic in which a conclusion may be justified on the basis of inconsistent premises should therefore give my opponent pause.²⁵

But I do not really have to rely on sympathies with the paraconsistency program in logic in order to make a case here. In the previous section, I argued that a physicist can be justified in rejecting the applicability of a mathematical technique if it differs too strikingly from the sorts of techniques originally intended when a mathematical concept was first introduced into the physicist's vocabulary. The fact that the physicist is justified in doing this follows from his agnosticism about the agreement problem. The physicist is ready to acknowledge the possibility that the agreement problem might have a negative answer, or that the way in which he currently interprets his physical concepts may not be compatible with an affirmative answer. Insofar as these are genuine possibilities for the physicist – and insofar as these possibilities do not especially bother him – he can justify conjecturing that certain unusual mathematical techniques ought not be applied in certain cases.

The final worry I wish to consider is that by allowing mathematically unrigorous arguments into physics, all sorts of absurdities become permissible. For instance, imagine that an experimental result is obtained in conflict with some theory. Let us further imagine that by deliberately not carrying the 1 in some particular addition problem involved in the theoretical calculation, agreement can be restored between theory and experiment. We might then imagine someone declaring that the experiment in question has shown that a new, unrigorous understanding of addition is needed in physics, and that the old theory is still unfalsified. As it stands, this is lunacy. But the worry is that if we allow *any* mathematically unrigorous arguments into physics, we will have no basis on which to condemn this sort of move.

The same worry can also be expressed a little differently. It might be suggested that the main reason certain physicists were so nonchalant about the nonrigorous mathematical methods of quantum field theory (especially in the days when renormalization was not as deeply understood as it is today) is that the theory had enjoyed truly extraordinary empirical success, and that this is really all physicists ought to care about – coming up with theories that accurately tell us what we will see in our particle accelerators. Thus, the use of non-

²⁴An important difference is that in the case of paraconsistent logic the restrictions on inference rules are usually based on the *logical form* of those rules – in the case of nonrigorous reasoning in physics, the restriction depends instead on the *content* of the propositions involved.

²⁵For a further discussion of reasoning with inconsistent scientific theories, see Smith [22] and Brown [3].

rigorous mathematical methods in physics is defended by pointing to the empirical success of relevant theories. The worry, however, is that if all that matters in physics is devising a calculus which can predict what we will see in our particle accelerators, then all sorts of ad-hoc physics is possible – such as the absurd example given in the previous paragraph.

My reply is that while some people might naively defend the use of unrigorous mathematics in physics by pointing to empirical successes and shrugging their shoulders, I have gone to great lengths to provide an *alternative* way of defending the use of unrigorous mathematical methods in physics. My defense certainly does not commit me to the view that all that matters to the physicist is that he come up with a calculus that can predict what will be seen in the particle accelerator. I fully agree that it is the job of physics to do more than this. The physicist does not only want to see his calculations work – he wants to understand *why* his calculations work. This requires, amongst other things, that he be able to view his calculations as articulating some set of physically intelligible processes that conspire to produce the empirical phenomenon in question. Our desire to *understand* the success of our calculations in this way is an important criterion of success for a theory in physics; and it is because the example given two paragraphs ago fails so drastically to meet this requirement that we correctly think of it as lunacy.

There is an obvious distinction to be drawn between an unrigorous physical theory for which we feel we understand why its calculations work (such as the path integral formulation of quantum mechanics), and an unrigorous physical theory for which we do not (such as the ‘pseudo-addition’ example given earlier). Insofar as it is possible to draw this distinction, the final worry of this section can be easily defused – the lunatic theories under consideration will not meet the criteria of a good physical theory for reasons ultimately unrelated to their mathematically unrigorous methodology.

The reader should be warned that none of the conclusions of this paper are intended to contradict the fact that attention to mathematical rigor can occasionally add a great deal to our understanding of physics. Furthermore, the careful and creative use of mathematics can sometimes lead to deep and lasting physical discoveries. But this fact can easily lead us to an inaccurate view of the goals and methods of physics, in which the role of rigor is elevated to an exaggerated status which does not truly belong to it. This sort of exaggeration is all I seek to avoid.

References

- [1] Beller, M., Quantum Dialogue : The Making of a Revolution, University of Chicago Press, 2001.
- [2] Bridgman, P., How Much Rigor is Possible in Physics? included as pp225-237 of The Axiomatic Methods with Special Reference to Geometry and Physics, Henkin, L., Suppes, P., Tarski, A. eds.,
- [3] Brown, B., How to be Realistic about Inconsistency in Science., Studies in the History and Philosophy of Science, 1990; vol 21, no 2, pp281-44.
- [4] Carnap, R., The Logical Structure of the World, University of California Press, 1967.
- [5] Cameron, R. H., A Family of Integrals Serving to Connect the Wiener and Feynman Integrals, Journal of Mathematics and Physics, 1960; 39, 126-140.
- [6] Dirac, P., The Principles of Quantum Mechanics, Oxford University Press, 1958.
- [7] Duhem, P., The Aim and Structure of Physical Theory, Princeton University Press, 1977.
- [8] Dicke, R., Wittke, J., Introduction to Quantum Mechanics, Addison-Wesley Publishing Company, 1960.
- [9] Feynman, R. and Hibbs, A., Quantum Mechanics and Path Integrals, McGraw-Hill, Inc., 1965.
- [10] Feynman, R. et al., The Feynman Lectures on Physics, Addison Wesley Longman, Inc, 1985.
- [11] Glimm, J., Jaffe, A., Quantum Mechanics - A Functional Integral Point of View, Springer Verlag, 1981.
- [12] Hempel, C., Aspects of Scientific Explanation and Other Essays in the Philosophy of Science, The Free Press, 1965.
- [13] Jammer, M., The Conceptual Development of Quantum Mechanics, McGraw Hill Book Company, 1966.
- [14] Kitcher, P., Mathematical Rigor – Who needs it?, Nous, vol 15, no 4, 1981, 469-493.

- [15] Landau, L. and Lifshitz, E., Statistical Physics, Oxford University Press, 1938.
- [16] Atiyah, M. et al., Responses to ‘Theoretical Mathematics: Toward a Cultural Synthesis of Mathematics and Theoretical Physics’, by A. Jaffe and F. Quinn, Bulletin of the American Mathematical Society, 1994; 30, 178-207
- [17] Malament, D., Is Newtonian Cosmology Really Inconsistent? Philosophy of Science, 1995; 62, 489-510.
- [18] Norton, J., A Paradox in Newtonian Gravitation Theory II, forthcoming.
- [19] Popper, K., The Logic of Scientific Discovery, Routledge, 1992.
- [20] Rivers, R. J., Path Integral Methods in Quantum Field Theory, Cambridge University Press, 1987.
- [21] Schwartz, L., Generalization de la notion de fonction, de derivation, de transformation de Fourier et applications mathematiques et physiques, Annales de l’Universite de Grenoble 21, 57-74, 1945.
- [22] Smith, J., Inconsistency and Scientific Reasoning, Studies in the History and Philosophy of Science, 1998; vol 19, no 4, pp429-445.
- [23] Steiner, M., Mathematical Rigor in Physics. In M. Detlefsen (ed.), Proof and Knowledge in Mathematics, Routledge, 1992.
- [24] Steiner, M., The Application of Mathematics to Natural Science, Journal of Philosophy, 1989; 86, 449-480.
- [25] von Neumann, J., Mathematical Foundations of Quantum Mechanics, Princeton University Press, 1996.
- [26] Wilson, M., On the Mathematics of Spilt Milk, from E. Grosholz and H. Breger (eds), The Growth of Mathematical Knowledge, 143-152.
- [27] Wilson, M., The Unreasonable Uncooperativeness of Mathematics in the Natural Sciences, The Monist, vol. 83, no 2, pp 296-314.